

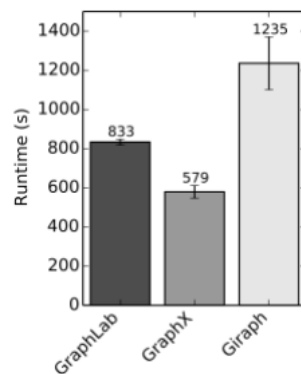
Weekly Report 2016.05.31-2017.06.04

Progress

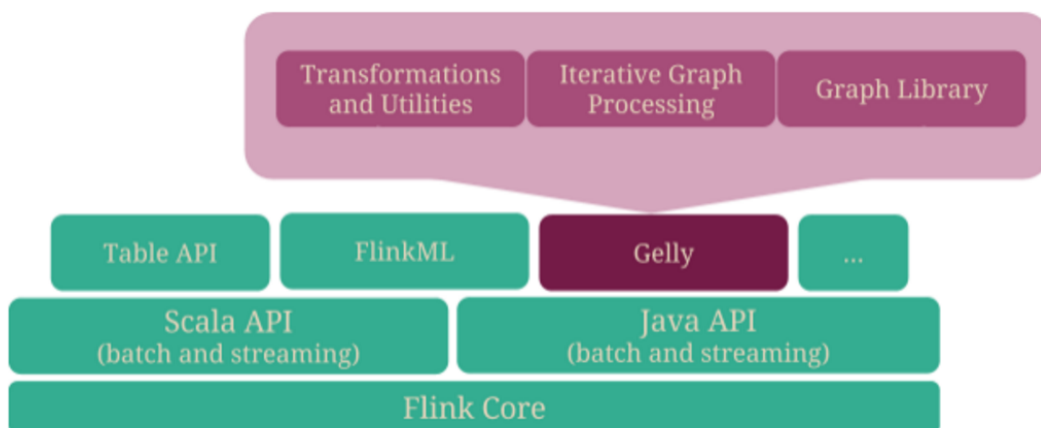
1. 大图可视化调研

本周调研了 flink, gelly 和 graphx 这三个 apache 项目。

首先, graphx 是 spark 提供的针对 graph 数据的 api, 可以支持图计算任务, 比如下图是计算一个具有 37 亿条边的图的 pagerank, 可以在 579 秒左右得到结果。



Flink 是一个与 spark 同级的计算平台, 主要针对的是 stream and batch data。而 gelly 则是与 graphx 同级的一个图计算库, 运行在 flink 上, 与 flink 的关系如下图:



Gelly 使用的是一种称为 vertex-centric 的编程模型, 可以运行一些迭代的图算法。

2. 可视化专著

我已经将图数据的大部分简介写完, 接下来撰写可视化技术部分, 这部分相对来

说比较好写，介绍一下常用方法即可。我找到了林涛的毕业论文，写的是当时 tievis 的论文，可以直接用到专著中。这样 rankflow 的工作用东明的毕业论文，rcanalyzer 的工作用嘉铨的毕业论文，而 tcptree 可以用斐然师兄的毕业论文，这样只要整合一下即可。